

## **Chapter 3**

**A. Polyamide Orientation and Affinity at the Sites 5'-GGTAG-3' and 5'-GATGG-3'; B. The Influence of (S)-2,4-diaminobutryic Acid on Orientation.**

*This research was completed in collaboration with Dr. Christian Melander.*

**Abstract**

Hairpin polyamides containing the  $\alpha$ -amino acids pyrrole, imidazole, and 3-hydroxypyrrole are reported to display a preference for a binding DNA in an orientation aligned such that the N $\rightarrow$ C polarity of each antiparallel DNA recognizing peptide strand of the polyamide is aligned with the 5 $\rightarrow$ 3' polarity of the phosphate backbone of the strand it recognizes ("forward" orientation). This chapter is devoted to the study of polyamides that tend to bind in an orientation *reversed* from the reported N $\rightarrow$ C, 5' $\rightarrow$ 3' orientation. From previous studies in our laboratory, this reversed orientation, whereby polyamides bind C $\rightarrow$ N, 5' $\rightarrow$ 3', is increasingly seen as a viable mode for polyamide recognition of the minor groove. In addition, we employ (*R*)-2,4-diaminobutyric acid as a monomer to enforce forward binding, and we describe for the first time the use of the acetylated (*S*)-2,4-diaminobutyric acid monomer to enforce reversed polyamide orientation to achieve specificity for reversed binding in the minor groove of DNA.

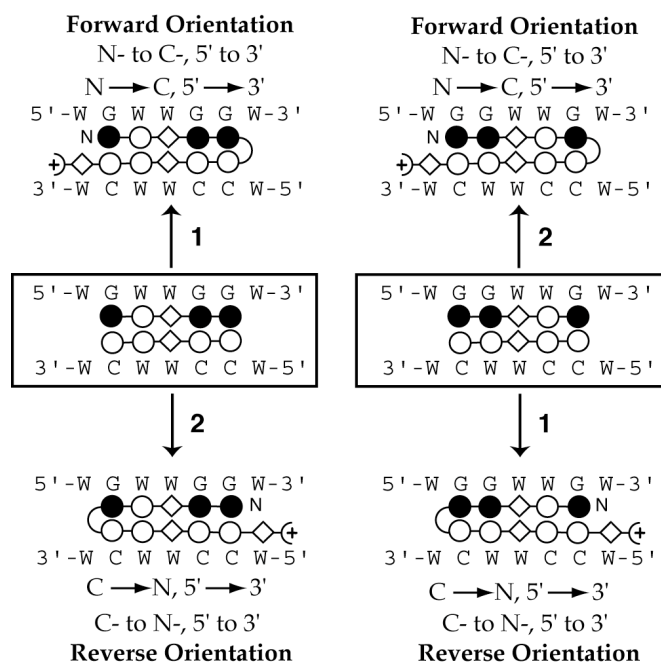
## Background

Early studies on the orientation preference of hairpin polyamides demonstrated that the polyamide preferentially align with the N-terminus of each antiparallel DNA strand recognition peptide aligned with the 5' end of the DNA strand recognized (N→C, 5'→3').<sup>1</sup> By DNase I footprinting, it was found that six-ring hairpin polyamides, containing only Im/Py base recognition elements, display a 16-fold binding bias for a binding site that allows the polyamide to align, forced by the pairing rules, in an N→C, 5'→3' orientation (as compared to a binding site whereby the pairing rules enforce reversed, C→N, 5'→3', binding). A growing body of experimental evidence in our group, however,<sup>2-7</sup> is suggesting that a reversed orientation, i.e., C→N, 5'→3', is wholly acceptable, and even energetically preferred, for some examples of polyamide recognition in the minor groove.

Our first interest in reversed binding comes from our study of polyamide recognition of the RCS of the HIV-1 LTR by 3- $\square$ -3 hairpin polyamides.<sup>2</sup> In this study, we observed a clear preference for reversed recognition of the sequence 5'-caGGCTCAGATct-3' by 3- $\square$ -3s in the minor groove. Building upon this observation, we noted that previous work by J. Trauger indicated a reversed binding mode for the polyamide ImIm- $\square$ -ImIm- $\square$ -PyPy- $\square$ -PyPy- $\square$ -Dp recognizing the sequence 5'-wwGGTGGww-3'.<sup>7</sup> More recently, a 1:1 motif has been shown to recognize DNA in a reversed orientation,<sup>3</sup> and NMR characterization of this binding mode suggests hydrogen bond donation from guanines in the minor groove to the hydrogen bond acceptors of imidazole confer the reversed orientation.<sup>4</sup> This suggests, as we and others have

observed, that the guanine rich strand of the DNA may control the polyamide orientation.<sup>4</sup>

As a test case for polyamide design, we chose the 2- $\square$ -2 motif, because of Trauger's observation and our own work with 3- $\square$ -3s, to explore polyamide orientation preferences. Building from Trauger's work with the core recognition sequence 5'-GGTGG-3', we chose the sequence 5'-GPuTPuG-3', where Pu = the purines A or G, and explored 5'-GGTAG-3' and 5'-GATGG-3'. As an example of the issues of polyamide design when reversed orientations are considered, we will go through the design of polyamides to target the sequence 5'-GATGG-3' (Figure 3.1). Placement of aromatic



**Figure 3.1** Ball-and-stick model for polyamide design. Open circles represent pyrrole, black circles represent imidazole, and open diamonds represent L-alanine. Left panel, considerations when designing polyamides to the sequence 5'-GWG-3'. Right panel, considerations when designing polyamides to the sequence 5'-GGW-3'.

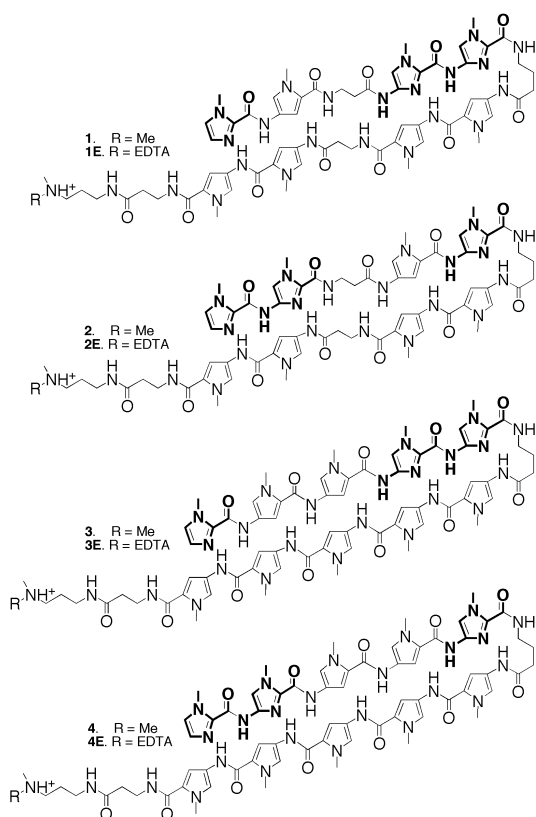
rings and L-alanine, without regard for the orientation of the hairpin, gives the following pairing rules read-out: Im/Py, Py/Py,  $\square/\square$ , Im/Py, Im/Py. Formally, the polyamide turn and tail would then be connected to allow the ligand to orient itself N  $\square$  C, 5'  $\square$  3'. This is what we would call a “forward” orientation (Figure 3.1, upper left panel). Reversing the placement of the  $\square$ -turn and  $\square$ -

Dp tail (Figure 3.1, lower left panel), results in a polyamide aligned C $\square$  N, 5'  $\square$  3', a “reverse” orientation, but still a pairing rules match for recognition of the sequence. A similar exercise can be carried out for recognition of the other sequence we studied, 5'-GGTAG-3' (Figure 3.1, right panels).

## Results and discussion

### The $\square/\square$ pair

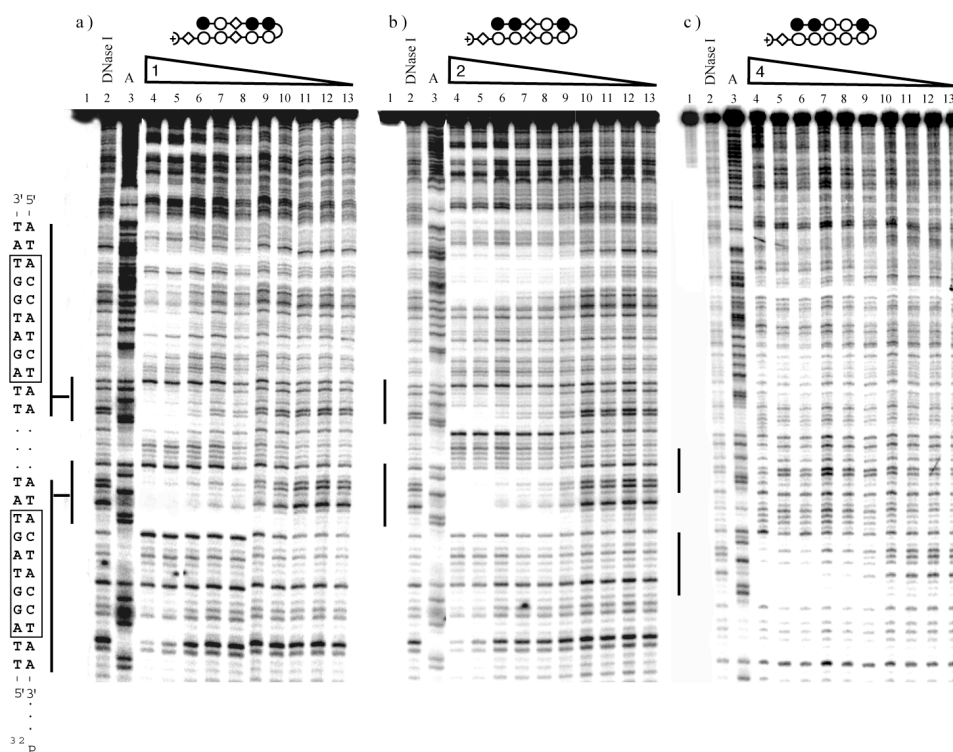
To probe for hairpin polyamides which may allow for, or prefer, a reversed orientation for DNA recognition, a pair of 2- $\square$ -2s (**1**, **2**), and their 10-ring relatives (**3**, **4**), were synthesized along with their affinity cleavage analogs (Figure 3.2).<sup>8</sup> The affinity for



**Figure 3.2** Polyamides **1-4**.

and orientation of these compounds in the minor groove was assessed by quantitative DNase I<sup>9</sup> and affinity cleavage<sup>10</sup> footprinting equilibration experiments against the radiolabeled restriction fragment from plasmid pV2 $\square$ 2. pV2 $\square$ 2 contains the binding sites 5'-GATGG-3' and 5'-GGTAG-3' allowing for assessment of binding orientation and affinity of polyamides **1** - **4** as described above. By the pairing rules, 5'-GATGG-3' presents itself as a forward oriented binding site for compounds **1** and **3** and a reverse oriented

binding site for compounds **2** and **4**. 5'-GGTAG-3', represents the opposite binding scenarios, i.e., it is a reverse site for **1/3** and a forward site for **2/4**.



**Figure 3.3** Quantitative DNase I footprint titration experiment with (a) ImPy-□-ImIm-□-PyPy-□-PyPy-□-Dp (**1**), (b) ImIm-□-PyIm-□-PyPy-□-PyPy-□-Dp (**2**), (c) ImImPyPyIm-□-PyPyPyPyPy-□□Dp (**4**) footprinted on the 5'-GGTAG-3' and 5'-GATGG-3' sites of the 3'-<sup>32</sup>P labeled restriction fragment from pVR2□2. Lane 1, intact DNA; lane 2, DNase I standard; polyamide concentration lanes 4-13, 50 nM, 20 nM, 1nM, 5 nM, 2 nM, 1 nM, 0.5 nM, 0.2 nM, 0.1 nM.

We observed by DNase I footprinting that polyamides **1** and **2** recognized each binding site with nanomolar to sub-nanomolar affinities such that **1** bound 5'-GATGG-3' with an affinity of  $8.2 \times 10^8 \text{ M}^{-1}$  and 5'-GGTAG-3' with an affinity of  $5.5 \times 10^9 \text{ M}^{-1}$ , while hairpin **2** bound 5'-GATGG-3' with an affinity of  $2.0 \times 10^9 \text{ M}^{-1}$  and 5'-GGTAG-3' with an affinity of  $4.1 \times 10^9 \text{ M}^{-1}$  (Figure 3.3, Table 3.1). DNase I footprinting analysis of polyamide **3** demonstrated the necessity of the □/□ pair to maintain nanomolar binding,

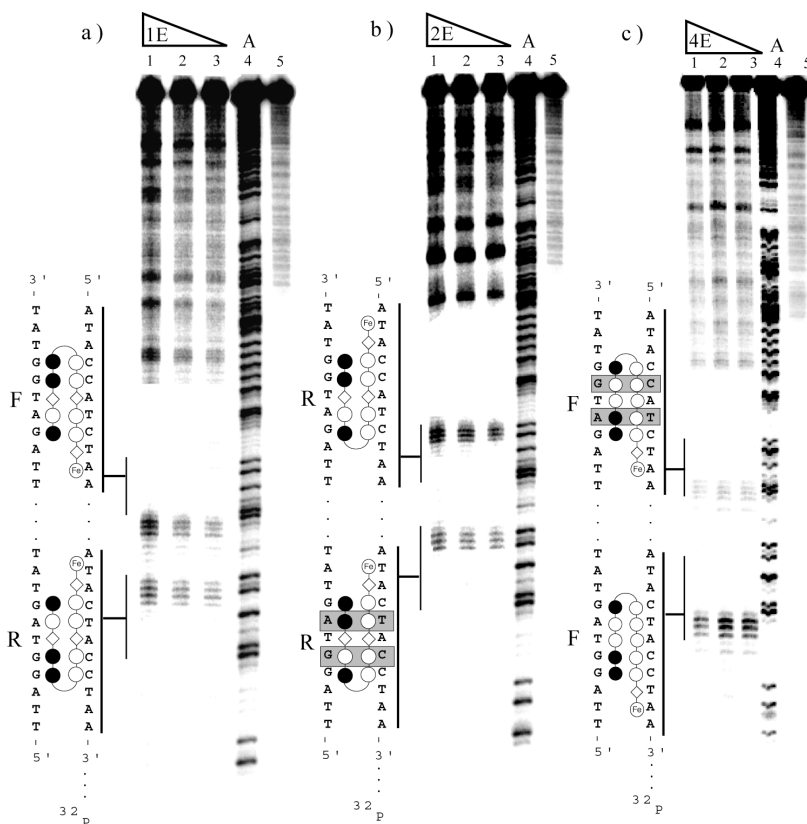
**Table 3.1** Equilibrium Association Constants ( $M^{-1}$ ) and Orientational Preference for Polyamides<sup>a-b</sup>

Polyamide	5'-tAGATGGTa-3'	Orientation	5'-tAGGTAGTa-3'	Orientation
1	$8.2 \times 10^8$ ( <b>0.8</b> )	Forward	$5.5 \times 10^9$ ( <b>0.3</b> )	Reverse
2	$2.0 \times 10^9$ ( <b>0.6</b> )	Reverse	$4.1 \times 10^9$ (0.3)	Reverse
3	$\leq 1 \times 10^7$	nd	$\leq 1 \times 10^7$	nd
4	$\leq 1 \times 10^7$	Forward	$5.0 \times 10^9$ ( <b>1.2</b> )	Forward

<sup>a</sup>Values reported are the mean values obtained from three DNase I titration experiments. <sup>b</sup>The assays were carried out at 22°C at pH 7.0 in the presence of 10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub>. nd = not determined (due to lack of binding). Match binding by pairing rules are indicated in boldface.

such that no binding at either site was observed up to polyamide concentrations of 50 nM (data not shown). Interestingly, hairpin **4** showed discrimination between the two binding sites and we observed a binding affinity of  $5.0 \times 10^9 M^{-1}$  for 5'-GATGG-3', while no binding up to 50 nM was observed at 5'-GGTAG-3' (Figure 3.3, Table 3.1).

We next determined the orientation for each hairpin polyamide using each ligand's affinity cleavage analog footprinted on the same radiolabeled fragment of pVR2□2. Analysis of the affinity cleavage pattern of polyamide **1E** (Figure 3.4) suggests both forward (5'-GATGG-3') and reversed binding (5'-GGTAG-3'), both sites bound with equal affinity (note: equienergetic refers to affinities determined from DNase I footprint titration experiments).<sup>11</sup> The cleavage pattern associated with hairpin **2E** (Figure 3.4) indicates that the polyamide obeys the pairing rules at the sequence 5'-GATGG-3' and binds in a reversed orientation. At the sequence 5'-GGTAG-3', **2E** prefers, again, to bind in a reversed orientation, but at this site it tolerates, with almost no energetic penalty (Table 3.1), a two base pair mismatch.



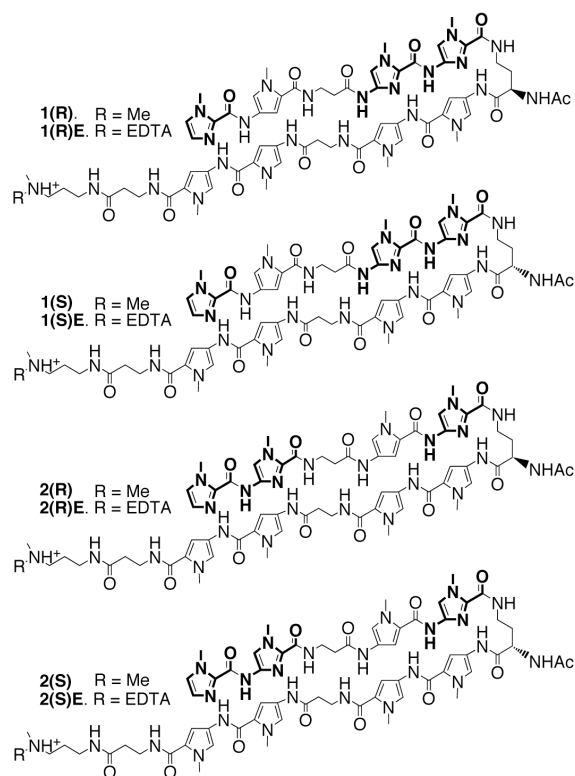
**Figure 3.4** Affinity cleavage experiments on the 3'-<sup>32</sup>P fill-in labeled restriction fragment from pVR2□2. 5'-GGTAG-3' and 5'-GATGG-3' sites with appropriately oriented ball-and-stick models of (a) ImPy-□-ImIm-□-PyPy-□-PyPy-□-EDTA (**1E**), (b) ImIm-□-PyIm-□-PyPy-□-PyPy-□-EDTA (**2E**), and (c) ImImPyPyIm-□-PyPyPyPyPy-□-EDTA (**4E**). Lanes 1-3, polyamide concentrations of 50 nM, 20 nM, 10 nM; lane 4, A sequencing lane; lane 5, intact DNA. All lanes contain 15 kcpm 5' kinase labeled, PCR amplified DNA, 25 mM-Tris acetate buffer (pH 7.0), 10 mM NaCl, and 100 mM/base pair calf thymus DNA.



binds 5'-GGTAG-3' as a match, as dictated by the pairing rules, and a double base pair mismatch at 5'-GATGG-3'. The subtle change of a pyrrole into a  $\beta$ -alanine completely switching the binding orientation observed for polyamide **2**.

### Stereochemical control of binding: Regaining specificity

In an effort to exercise our understanding of molecular interactions that affect polyamide•DNA recognition, we investigated stereochemistry as a means to control



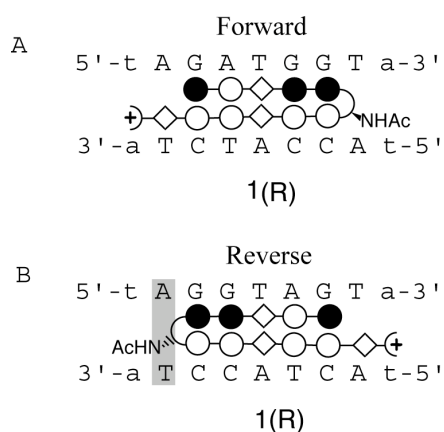
**Figure 3.5** Polyamides **1(R)**-**2(S)**.

binding orientation within the 2- $\beta$ -2 scaffold. Earlier studies demonstrated that incorporation of the (R)-2,4-diaminobutyric acid subunit,<sup>12</sup> a chiral analog of the  $\beta$ -turn, enforced the preference for forward (N  $\rightarrow$  C, 5'  $\rightarrow$  3') binding of a six-ring, Py/Im hairpin in the minor groove. The (S)-2,4-diaminobutyric acid unit was also been investigated as a chiral replacement for the  $\beta$ -turn in the forward binding six-ring context,<sup>12</sup> but a substantial reduction in the observed equilibrium

constant was observed. This was attributed to a steric clash between the primary  $\beta$ -amine on the turn and the floor of the minor groove. It was unclear if incorporation of

the S stereochemistry could be employed, alternatively, to enforce polyamide binding in a C $\square$ N, 5' $\square$ 3 orientation.

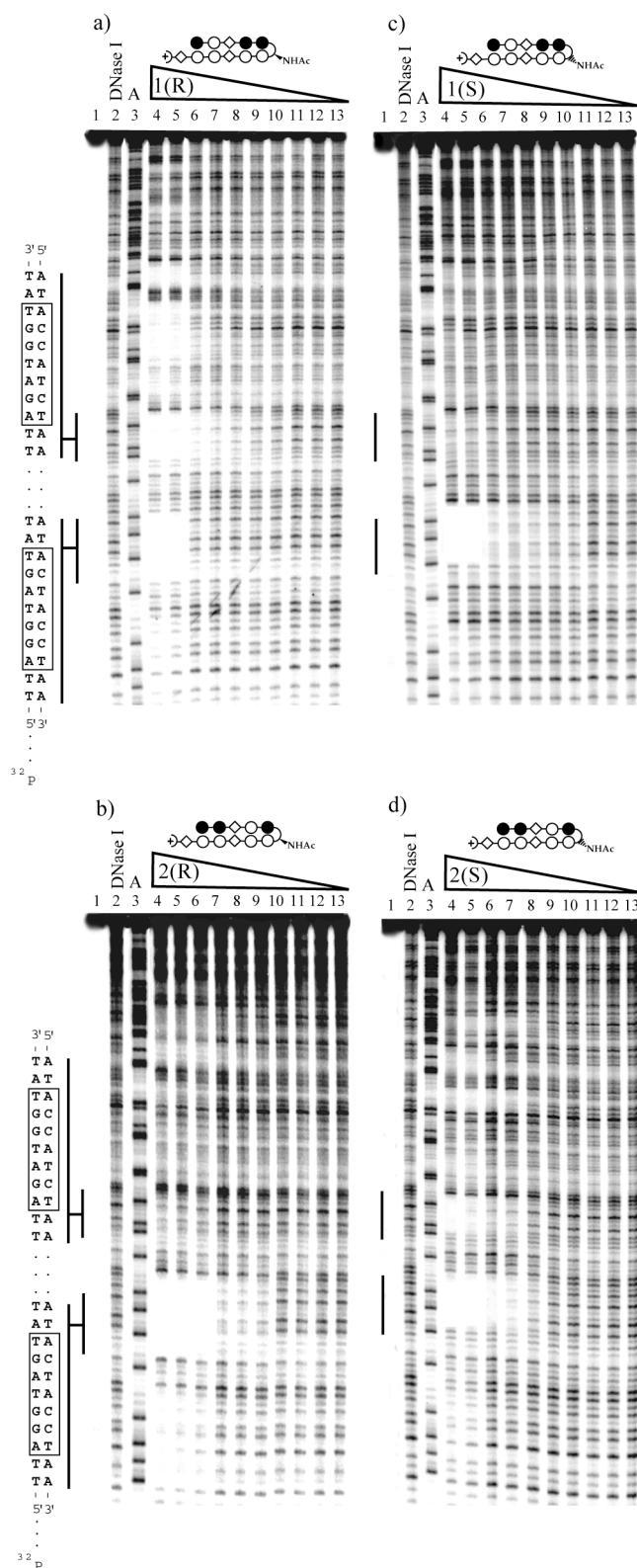
We addressed this issue of stereocontrol by employing both the R and S isomer of 2,4-diaminobutyric acid to explore its effect on the binding orientation of polyamides **1** and **2**. Chiral analogs of **1** and **2** were synthesized (**1R**, **1S**, **2R**, **2S**), as well as their affinity cleavage analogs (**1(R)E**, **1(S)E**, **2(R)E**, **2(S)E**) (Figure 3.5).<sup>8,12</sup> The primary  $\square$ -amine of the chiral diaminobutyric acid turns were acetylated to conserve overall charge in comparison to hairpins **1** and **2**, as well as to eliminate any energetic binding benefits



**Figure 3.6** A. **1R** binding in a forward orientation to 5'-GATGG-3', the  $\square$ -N-acetyl projects away from the minor groove and imposes no steric impediment to binding. B. **1R** binding 5'-GGTAG-3' in a reverse orientation, the  $\square$ -N-acetyl projects into the minor groove, incurring an energetic penalty to disfavor binding.

gained by placement of the second charge.

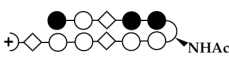
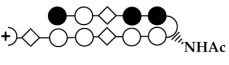
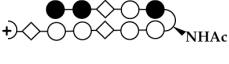

Earlier observations noted that employment of the (R)-2,4-diaminobutyric acid turn resulted in a 10-fold increase in binding affinity, while acetylating the primary amine (as in **1R** - **2S**) produced ligands with binding affinities equal to those of the parent hairpin polyamides containing the underivatized  $\square$ -turn.<sup>12</sup> The polyamides in Figure 3.5 were investigated by affinity cleavage and DNase I footprint titration against the two binding sites 5'-GGTAG-3' and 5'-GATGG-3' on the radiolabeled fragment of pVR2 $\square$ 2.



**Figure 3.7** (a) **1R**, (b) **2R**, (c) **1S**, (d) **2S** footprinted on the 5'-GGTAG-3' and 5'-GATGG-3' sites of the 5'-<sup>32</sup>P labeled restriction fragment from pVR2[2]. Lane 1, intact DNA; lane 2, DNase I standard; polyamide concentration lanes 4-13, 50 nM, 20 nM, 1nM, 5 nM, 2 nM, 1 nM, 0.5 nM, 0.2 nM, 0.1 nM.

If we consider polyamide **1R**, an analysis of the predicted interaction between the acetylated  $\beta$ -amine and the floor of the minor groove, with regard to the ligand orientation, can be made (Figure 3.6). When **1R** is bound in a forward orientation, the acetylated  $\beta$ -amino group of the turn projects itself away from the floor of the minor groove, and there is no energetic penalty for binding. However, if **1R** were to bind in a reversed orientation, the  $\beta$ -N-acetyl moiety would protrude into the floor of the minor groove and sterically clash with the floor of the minor groove. This would, presumably, incur an energetic penalty and disfavor this binding mode. For all of the stereochemical derivatives in Figure 3.5, a similar scenario can be envisioned.

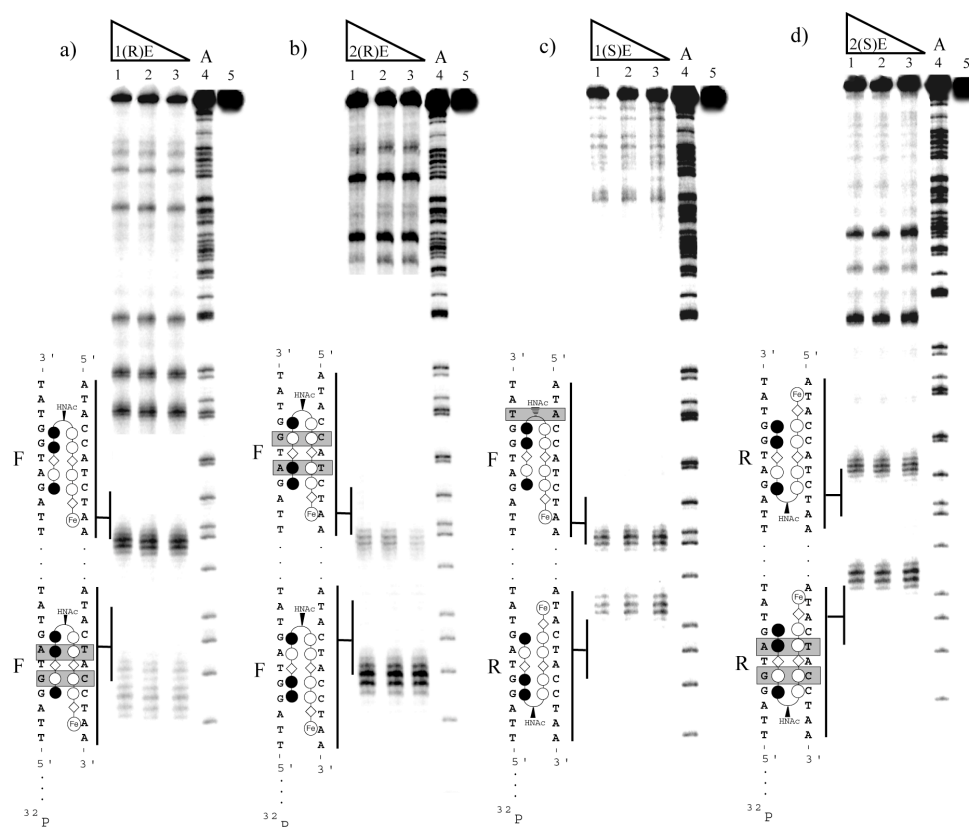
**Table 3.2** Equilibrium Association Constants ( $M^{-1}$ ) and Orientational Preference for Chiral Polyamides<sup>a-b</sup>

Polyamide	5'-tAGATGGTa-3'	Orientation	5'-tAGGTAGTa-3'	Orientation
<b>1(R)</b> 	<b><math>8.6 \times 10^9</math></b> ( <b>0.8</b> )	Forward	$6.8 \times 10^8$ (0.3)	Forward
<b>1(S)</b> 	$1.9 \times 10^7$ (0.8)	Forward	<b><math>1.2 \times 10^9</math></b> ( <b>0.3</b> )	Reverse
<b>2(R)</b> 	$2.7 \times 10^7$ (0.6)	Forward	<b><math>1.0 \times 10^9</math></b> ( <b>0.3</b> )	Forward
<b>2(S)</b> 	<b><math>3.0 \times 10^8</math></b> ( <b>1.2</b> )	Reverse	$1.3 \times 10^9$ (0.3)	Reverse

<sup>a</sup>Values reported are the mean values obtained from three DNase I titration experiments. <sup>b</sup>The assays were carried out at 22°C at pH 7.0 in the presence of 10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub>. Match binding by "pairing rules" are indicated in boldface.

Quantitative DNase I footprinting of **1R** against the restriction fragment of pV2[2] gave observed equilibrium constants of  $8.6 \times 10^9 M^{-1}$  for 5'-GATGG-3' and  $6.8 \times 10^8 M^{-1}$  for 5'-GGTAG-3', while DNase I footprinting of **1S** produced the opposite trend and we observed equilibrium association constants of  $1.9 \times 10^7 M^{-1}$  for 5'-GATGG-3' and  $1.2 \times 10^9 M^{-1}$  for 5'-GGTAG-3' (Figure 3.7, Table 3.2). We were gratified to note that

while the parent compound, polyamide **1**, showed only modest selectivity (a 6-fold preference for 5'-GGTAG-3' over 5'-GATGG-3'), incorporation of the R isomer reversed this trend and hairpin **1R** bound its match site 5'-GGTAG-3' with a 12-fold preference over 5'-GATGG-3' and retention of good affinity. Polyamide **1S** also displayed subnanomolar affinity at its reversed match site, 5'-GGTAG-3', while augmenting the



**Figure 3.8** Affinity cleavage experiments on the 5'-<sup>32</sup>P labeled restriction fragment from pVR2 $\Delta$ 2. 5'-GGTAG-3' and 5'-GATGG-3' sites with appropriately oriented ball-and-stick models of (a) ImPy-ImIm-(R)<sup>AcNH</sup>-PyPy-PyPy-EDTA (**1(R)E**), (b) ImIm-PyIm-(R)<sup>AcNH</sup>-PyPy-PyPy-EDTA (**2(R)E**), (c) ImPy-ImIm-(S)<sup>AcNH</sup>-PyPy-PyPy-EDTA (**1(S)E**), (d) ImIm-PyIm-(S)<sup>AcNH</sup>-PyPy-PyPy-EDTA (**2(S)E**). Lanes 1-3, polyamide concentrations of 50 nM, 20 nM, 10 nM; lane 4, A sequencing lane; lane 5, intact DNA. All lanes contain 15 kcpm 5' kinase labeled, PCR amplified DNA, 25 mM-Tris acetate buffer (pH 7.0), 10 mM NaCl, and 100  $\mu$ M/base pair calf thymus DNA.

preference for 5'-GGTAG-3' over 5'-GATGG-3' from 6- to 63-fold. DNase I footprinting analysis of polyamide **2R** resulted in equilibrium association constants of  $2.7 \times 10^7 \text{ M}^{-1}$  for 5'-GATGG-3' and  $1.0 \times 10^9 \text{ M}^{-1}$  for 5'-GGTAG-3' (Figure 3.7, Table 3.2). This represents, in comparison to polyamide **2**, and increase in preference from 0.5- to a 37-fold for binding 5'-GGTAG-3' with respect to 5'-GATGG-3'. In accordance with the observation that parent hairpin polyamide **2** binds both sites in a reversed orientation (Figure 3.4), it was anticipated that incorporation of the S enantiomer of 2,4-diaminobutyric acid would not exert a substantial effect on discrimination between the binding sites 5'-GATGG-3' and 5'-GGTAG-3'. DNase I footprinting of **2S** indicated that this was indeed the case, and we observe equilibrium association constants of  $3.0 \times 10^8 \text{ M}^{-1}$  and  $1.3 \times 10^9 \text{ M}^{-1}$  for binding of 5'-GATGG-3' and 5'-GGTAG-3' respectively (a 4.3-fold discrimination versus 2.05-fold for the parent).

Analysis of the affinity cleavage data for hairpins **1(R)E** – **2(S)E** indicates the effect of a stereochemically derivatized turn upon orientation (Figure 3.8). In both **1R** and **2R**, we observe binding solely in a forward orientation at both sites. This indicates that stereocontrol is the energetically overriding factor in binding orientation for **1R** as we have disfavored reversed binding at the site 5'-GGTAG-3' that was observed for the parent, **1**. **1** also tolerates a two base-pair mismatch to bind this site. This is especially interesting, in light of the observation that polyamide **2** binds both 5'-GATGG-3' and 5'-GGTAG-3' in a reverse orientation. Thus, we stereochemically influenced the tendency for parent polyamide **2** to bind the site 5'-GGTAG-3' in a reversed orientation. Incorporation of the S enantiomer in polyamide **2**, i.e., **2S**, had no affect upon orientation as both sites were bound by parent **2** in a reversed orientation. The S enantiomer serves

to reinforce this reversed binding mode preference. Hairpin **1S** bound 5'-GGTAG-3' in a reverse orientation and 5'-GATGG-3' in a forward orientation, thus preserving the orientation observed for polyamide **1** in spite of a steric clash between the  $\beta$ -N-acetyl moiety and the floor of the minor groove. Although this observation may at first seem discouraging, it is not. Although we did not achieve a binary reversal of polyamide orientation in accordance with stereochemical predictions, incorporating stereochemistry at the  $\beta$  position of the  $\beta,\beta$ -diamino butyric acid monomer did have the designed effect. We achieved better energetic discrimination between the two binding sites than was observed for the parent polyamide, i.e., **1S** discriminates 63-fold between the sites 5'-GGTAG-3' and 5'-GATGG-3' compared to **1** which only shows 6.7-fold discrimination. In the scenarios presented here, we are clearly probing a fine energetic balance between satisfying the hydrogen bonds of the polyamide in the minor groove, opposing or reinforcing the inherent orientation preference of the polyamide, and avoiding steric clashes between the moieties at the  $\beta$  position of the  $\beta,\beta$ -diaminobutyric acid monomer and the floor of the minor groove.

## Conclusions

The design of polyamides to recognize and bind DNA sequences of increasing size has necessitated the introduction of flexible units. These residues allow the polyamide to correctly register with the hydrogen bond donor and acceptor patterns displayed in the minor groove of the DNA helix, but it may also confer variability to polyamide orientation. The importance of  $\beta$ -alanine has arisen from observations that i) G•C base pairs recognized in the internal region of the polyamide sequence are

traditionally more difficult to properly register and require an adjacent flexible unit to allow optimal polyamide/DNA interactions<sup>13</sup> and ii) although polyamides match the rise per residue of the helix, they become overwound as they increase in size.<sup>14</sup> The introduction of the  $\square/\square$  pair has overcome this pitfall, and provided ligands with increased flexibility and the ability to bind larger DNA sequences, while largely maintaining high affinities and good sequence specificity.<sup>13</sup>

Unexpectedly, the introduction of the  $\square/\square$  pair in compounds that target sequences with a predominantly G-rich strand has interestingly added to the previously reported polyamide alignment preference of N $\square$  C, 5' $\square$  3' determined for six-ring hairpins.<sup>1</sup> We have shown that introduction of the  $\square/\square$  pair, as a replacement for the Py/Py pair, provides ligands that bind DNA with high affinity in both a forward and reversed orientation.<sup>5</sup> Substitution of this  $\square/\square$  pair with a Py/Py pair appears to strictly enforce a N $\square$  C, 5' $\square$  3' orientation preference (**4**), in accordance with the orientation reported for all-ring six-ring hairpins, but binding of a match site is not always observed (**3**). Clearly, a Py/Py substitution for a  $\square/\square$  pair in 10-ring polyamides complicates the process of predicting *a priori* a polyamide's orientation.

The mechanism by which the  $\square/\square$  pair allows or even prefers a C- to N-, 5'- to 3'- polyamide•DNA recognition motif is unclear, but it may be related to hydrogen bond directionality between imidazole hydrogen bond acceptors in the polyamide and guanine hydrogen bond donors in the minor groove.<sup>4</sup> Regardless of the reason for reversed orientation, the additions to polyamide design discussed here are 2-fold. First, when incorporation of a  $\square/\square$  pair is required to maintain binding affinity, recognition elements should first be placed at the DNA sequence of interest regardless of orientation (as in



Figure 3.1). Second, use of N- $\alpha$ -acetyl (R or S) 2,4-diaminobutyric acid can be employed to augment specificity, dependent upon the polyamide orientation, by energetically favoring either the forward (R) or reversed (S) binding mode.

## Experimental

Dicyclohexylcarbodiimide (DCC), Hydroxybenzotriazole (HOBt), 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexa-fluorophosphate (HBTU), Boc- $\alpha$ -alanine ( $\alpha$ ) and 0.26 mmol/gram Boc- $\alpha$ -alanine-(4-carboxamidomethyl)-benzyl-ester-copoly(styrene-divinylbenzene) resin (Boc- $\alpha$ -Pam-Resin) was purchased from Peptides International (0.26 mmol/gram) (*R*)-2-Fmoc-4-Boc-diaminobutyric acid and (*S*)-2-Fmoc-4-Boc-diaminobutyric acid were from Bachem. *N,N*-diisopropylethylamine (DIEA), *N,N*-dimethylformamide (DMF), *N*-methylpyrrolidone (NMP), DMSO/NMP, and Acetic anhydride (Ac<sub>2</sub>O) were from Applied Biosystems. Dichloromethane (DCM) and triethylamine (TEA) were reagent grade from EM, thiophenol (PhSH), 3,3'-diamino-*N*-methyldipropylamine and dimethylaminopropylamine (Dp) were from Aldrich. Trifluoroacetic acid (TFA) Biograde was from Halocarbon. All reagents were used without further purification.

Quik-Sep polypropylene disposable filters were purchased from Fisher. A shaker for manual solid phase synthesis was obtained from VWR. Screw-cap glass peptide synthesis reaction vessels (5 mL and 20 mL) with a #2 sintered glass frit were made as described by Kent.<sup>15</sup> UV spectra were measured in water on a Hewlett-Packard Model 8452A diode array spectrophotometer. Matrix-assisted, laser desorption/ionization time of flight mass spectrometry (MALDI-TOF) was performed at the Protein and Peptide Microanalytical Facility at the California Institute of Technology. HPLC analysis was performed on either a HP 1090M analytical HPLC or a Beckman Gold system using a RAINEN C<sub>18</sub>, Microsorb MV, 5  $\mu$ m, 300 x 4.6 mm reversed phase column in 0.1%

(wt/v) TFA with acetonitrile as eluent and a flow rate of 1.0 mL/min, gradient elution 1.25% acetonitrile/min. Preparatory reverse phase HPLC was performed on a Beckman HPLC with a Waters DeltaPak 25 x 100 mm, 100  $\mu$ m C18 column equipped with a guard, 0.1% (wt/v) TFA, 0.25% acetonitrile/min. 18M $\Omega$  water was obtained from a Millipore MilliQ water purification system, and all buffers were 0.2  $\mu$ m filtered.

Enzymes were purchased from Boehringer-Mannheim and used with their supplied buffers. Deoxyadenosine and thymidine 5'-[ $\gamma$ - $^{32}$ P] triphosphates were obtained from Amersham, and deoxyadenosine 5'-[ $\gamma$ - $^{32}$ P]triphosphate was purchased from I.C.N. Sonicated, deproteinized calf thymus DNA was acquired from Pharmacia. RNase free water was obtained from USB and used for all footprinting reactions. All other reagents and materials were used as received. All DNA manipulations were performed according to standard protocols.<sup>16</sup>

### **Polyamide synthesis**

Reagents and protocols for manual, solid-phase polyamide synthesis were as described.<sup>8,12</sup> Polyamides were liberated from resin with either dimethylamino-propylamine (Dp) or 3,3'-diamino-N-methyldipropylamine at 55 °C for 16 hrs and purified by reversed-phase HPLC with Waters DeltaPak 25 x 100 mm, 100  $\mu$ m C18 column equipped with a guard, and eluted with 0.1% (wt/v) TFA, 8.0 mL/min, 0.25% acetonitrile/min. Affinity cleavage analogs were synthesized as previously described.<sup>12</sup> Extinction coefficients were calculated based on  $\epsilon$  = 8600 M<sup>-1</sup>cm<sup>-1</sup>/ring at 310 nm.<sup>17</sup> Compound purity and identity was verified for all compounds by analytical HPLC and MALDI/TOF MS. For **3** (monoisotopic) 1452.64 calc'd for C<sub>68</sub>H<sub>80</sub>N<sub>26</sub>O<sub>12</sub><sup>+</sup>, 1453.52

found for [M + H]; for **4** 1452.64 calc'd for  $C_{68}H_{80}N_{26}O_{12}^+$ , 1453.78 found for [M + H]; for **1** 1350.62 calc'd for  $C_{62}H_{78}N_{24}O_{12}^+$ , 1351.44 found for [M + H]; for **2** 1350.62 calc'd for  $C_{62}H_{78}N_{24}O_{12}^+$ , 1351.72 found for [M + H]; for **5R**, **5S**, **6R**, **6S** 1407.64 calc'd for  $C_{64}H_{81}N_{25}O_{13}^+$  and ([M + H] values) 1408.24 found for **5R**, 1408.1 found for **5S**, 1408.68 found for **6R**, and 1408.52 found for **6S**.

### Construction of plasmids pVR2□2

*Bam*HI/*Hind*III restriction sites were chosen for ligation of synthetic double-stranded DNA inserts into the polylinker cloning site of pUC-19. Inserts for pVR2□2 were all provided as dry single stranded oligos as synthesized by the Biopolymer Facility at the California Institute of Technology. The single stranded oligos are as follows: pVR2□2 5'-GATCCGGCCCTTAGATAGTATGGCCACGTGGTAGTGGCCACGTGATGGTGGCCCA-3', 5'-AGCTTGGGCCATACCATCTAACGTGGGCCATACTACCTAACGTGGGCCATACTATCTAAGGGCCG-3'. Double stranded DNA inserts were ligated in pUC-19 as per the protocol provided in the Roche Rapid DNA Ligation Kit. The constructed plasmids were then transformed into Promega JM109 *E. coli* Subcloning Efficiency Competent Cells. Successful ligations and transformations were chosen on the basis of blue/white screening in the presence of IPTG and X-Gal. Colonies were chosen, grown, and the plasmids harvested using the Promega Wizard® *Plus* Midiprep purification system. Plasmids were sequenced at the Nucleic Acids Sequencing Facility at The California Institute of Technology. Typical yields for purification of 75 mL of *E. coli* cells grown for 24 hrs at 37 °C in LB medium are 0.347 µg/µL.

### Preparation of 3'- and 5'-end-labeled restriction fragments

For 3' label, the appropriate plasmid was linearized with *EcoRI* and *PvuII*, then treated with the Sequenase enzyme, deoxyadenosine 5'-[ $\gamma$ -<sup>32</sup>P]triphosphate and thymidine 5'-[ $\gamma$ -<sup>32</sup>P]triphosphate for 3' labeling. For 5'-labeling: two 20 base pair primer oligonucleotides, 5'- A A T T C G A G C T C G G T A C C C G G -3' (forward) and 5'- C T G G C A C G A C A G G T T T C C C G -3' (reverse) were constructed to complement the pUC19 *EcoRI* and *PvuII* sites, respectively, such that amplification by PCR would mimic the long, 3'-filled, pUC19 *EcoRI/PvuII* restriction fragment. The forward primer was radiolabelled using  $\gamma$ -<sup>32</sup>P-dATP and polynucleotide kinase, and the appropriate region amplified using PCR. The labeled fragment (3' or 5') was loaded onto a 6% non-denaturing polyacrylamide gel, and the desired band was visualized by autoradiography and isolated. Chemical sequencing reactions were performed according to published methods.<sup>18</sup>

### Affinity cleaving<sup>10</sup>

All reactions were carried out in a volume of 40  $\mu$ L. A polyamide stock solution or water (for reference lanes) was added to an assay buffer where the final concentrations were: 25 mM Tris-acetate buffer (pH 7.0), 20 mM NaCl, 100  $\mu$ M/base pair calf thymus DNA, and 20 kcpm 3'- or 5'-radiolabeled DNA. The solutions were allowed to equilibrate for 8 hrs. A fresh solution of ferrous ammonium sulfate ( $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}$ ) (10  $\mu$ M) was added to the equilibrated DNA, and the reactions were allowed to equilibrate for 15 mins. Cleavage was initiated by the addition of

dithiothreitol (10 mM) and allowed to proceed for 30 min. Reactions were stopped by ethanol precipitation, resuspended in 100 mM tris-borate-EDTA/80% formamide loading buffer, denatured at 85 °C for 6 min, and the entire sample was immediately loaded onto an 8% denaturing polyacrylamide gel (5% cross-link, 7 M urea) at 2000 V.

### **DNase I footprinting<sup>9</sup>**

All reactions were carried out in a volume of 400  $\mu$ L. We note explicitly that no carrier DNA was used in these reactions until after DNase I cleavage. A polyamide stock solution or water (for reference lanes) was added to an assay buffer where the final concentrations were: 10 mM Tris•HCl buffer (pH 7.0), 10 mM KCl, 10 mM MgCl<sub>2</sub>, 5 mM CaCl<sub>2</sub>, and 30 kcpm 3'-radiolabeled DNA. The solutions were allowed to equilibrate for a minimum of 12 hrs at 22 °C. Cleavage was initiated by the addition of 10  $\mu$ L of a DNase I stock solution (diluted with 1 mM DTT to give a stock concentration of 1.875 U/mL) and was allowed to proceed for 7 min at 22 °C. The reactions were stopped by adding 50 mL of a solution containing 2.25 M NaCl, 150 mM EDTA, 0.6 mg/mL glycogen, and 30 mM base-pair calf thymus DNA, and then ethanol precipitated. The cleavage products were resuspended in 100 mM tris-borate-EDTA/80% formamide loading buffer, denatured at 85 °C for 6 min, and immediately loaded onto an 8% denaturing polyacrylamide gel (5% cross-link, 7 M urea) at 2000 V for 1 hr. The gels were dried under vacuum at 80 °C, then quantitated using storage phosphor technology.

Equilibrium association constants were determined as previously described.<sup>9</sup> The data were analyzed by performing volume integrations of the appropriate sites and a

reference site. The apparent DNA target site saturation,  $\square_{app}$ , was calculated for each concentration of polyamide using the following equation:

$$\square_{app} = 1 - \frac{I_{tot}/I_{ref}}{I_{tot}^{\circ}/I_{ref}^{\circ}} \quad (1)$$

where  $I_{tot}$  and  $I_{ref}$  are the integrated volumes of the target and reference sites, respectively, and  $I_{tot}^{\circ}$  and  $I_{ref}^{\circ}$  correspond to those values for a DNase I control lane to which no polyamide has been added. The  $([L]_{tot}, \square_{app})$  data points were fit to a Langmuir binding isotherm (eq 2,  $n=1$  for polyamides **1-3**,  $n=2$  for polyamides **4** and **5**) by minimizing the difference between  $\square_{app}$  and  $\square_{fit}$ , using the modified Hill equation:

$$\square_{fit} = \square_{min} + (\square_{max} - \square_{min}) \frac{K_a^n [L]_{tot}^n}{1 + K_a^n [L]_{tot}^n} \quad (2)$$

where  $[L]_{tot}$  corresponds to the total polyamide concentration,  $K_a$  corresponds to the equilibrium association constant, and  $\square_{min}$  and  $\square_{max}$  represent the experimentally determined site saturation values when the site is unoccupied or saturated, respectively. Data were fit using a nonlinear least-squares fitting procedure of KaleidaGraph software with  $K_a$ ,  $\square_{max}$ , and  $\square_{min}$  as the adjustable parameters. All acceptable fits had a correlation coefficient of  $R > 0.98$ . At least three sets of acceptable data were used in determining each association constant. All lanes from each gel were used unless visual inspection

revealed a data point to be obviously flawed relative to neighboring points. The data were normalized using the following equation:

$$\square_{\text{norm}} = \frac{\square_{\text{app}} - \square_{\text{min}}}{\square_{\text{max}} - \square_{\text{min}}} \quad (3)$$

### **Quantitation by storage phosphor technology autoradiography**

Photostimulable storage phosphorimaging plates (Kodak Storage Phosphor Screen S0230 obtained from Molecular Dynamics) were pressed flat against gel samples and exposed in the dark at 22 °C for 12-20 hr. A Molecular Dynamics 400S PhosphorImager was used to obtain all data from the storage screens. The data were analyzed by performing volume integrations of all bands using the ImageQuant v. 3.2.



## References

1. White, S.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1997**, *119*, 8756.
2. Coull, J.J.; He, G.; Melander, C.; Rucker, V.C.; Dervan, P.B., Margolis, D.M. *J. Virol.* **2002**, *76*, 12349.
3. Urbach, A. R.; Dervan, P.B. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4343.
4. Urbach, A. R.; Love, J. J.; Ross, S. A.; Dervan, P. B. *J. Mol. Biol.* **2002**, *320*, 55.
5. Chapter 2A of this thesis.
6. Hawkins, C. A.; de Clairac, R. N.; Dominey, R. N.; Baird, E. E.; White, S.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **2000**, *122*, 5235.
7. Trauger, J. W. Thesis, California Institute of Technology.
8. Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141.
9. Trauger, J.W.; Dervan, P.B. *Methods Enz.* **2001**, *340*, 450.
10. Taylor, J. S.; Schultz, P. G.; Dervan, P. B. *Tetrahedron* **1984**, *40*, 457.
11. Affinity cleavage is only a measure of a ligand's orientation. Association constants are only determined by DNase I footprint titration experiments and analysis.
12. Herman, D. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, *120*, 1382.
13. Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am Chem. Soc.* **1998**, *120*, 6219.
14. Kelly, J. J.; Baird, E. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 6981.
15. Kent, S. B. H. *Annu. Rev. Biochem.* **1988**, *57*, 957.
16. Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1989.
17. Pilch, D. S.; Pokar, N. A.; Gelfand, C. A.; Law, S. M.; Breslauer, K. J.; Baird, E. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8306.
18. Iverson, B. L.; Dervan, P. B. *Nucl. Acids Res.* **1987**, *15*, 7823.